

# Descriptive Statistics in Clinical Trials

Shu-Ping Chen

2020/10/04

## Descriptive Statistics in Clinical Trials

### A. Leprosy Clinical Trial

Participants in a clinical trial were randomized to either of two antibiotics (denoted treatment drug A and B) or to a placebo (denoted treatment drug C). Prior to receiving treatment, baseline data on the number of leprosy bacilli at six sites of the body where the bacilli tend to congregate were recorded for each patient. After several months of treatment, the number of leprosy bacilli at six sites of the body were recorded a second time. The outcome variable is the total count of the number of leprosy bacilli at the six sites. In this study, the question of the main scientific interest is whether treatment with antibiotics (drugs A and B) reduces the abundance of leprosy bacilli at the six sites of the body when compared to placebo (drug C). Data are in files CTleprosyWide.csv and CTleprosyLong.csv.


### Introduction


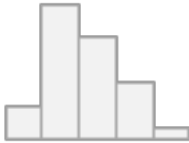
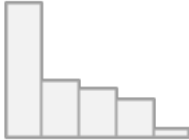
This trial makes the patients randomly distributed to Rifampin, Dapsone, and Placebo group. The group A, B, and C is Rifampin, Dapsone, and Placebo group, respectively. The trial records whether the number of leprosy bacilli at six sites of the patients' body before the treatment and after treatment for few months. Through Table 1, there isn't any missing values of variables of data, and the number of patients in each group is the same. The number of leprosy bacilli before the treatment is normally distributed; the number of leprosy bacilli after the treatment is skewed to the right. Hence, we want to observe whether treatment with antibiotics (Rifampin and Dapsone) reduces the abundance of leprosy bacilli at the six sites of the body when compared to placebo (drug C).

Table 1: Information of data

**Dimensions:** 30 x 4

**Duplicates:** 0

Variable	Stats / Values	Freqs (% of Valid)	Graph	Missing
id [integer]	Mean (sd) : 15.5 (8.8) min < med < max: 1 < 15.5 < 30 IQR (CV) : 14.5 (0.6)	30 distinct values (Integer sequence)		0 (0%)

treat	1. A	10 (33.3%)		0
[character]	2. B	10 (33.3%)		(0%)
	3. C	10 (33.3%)		
baseline	Mean (sd) : 10.7 (4.8)	16 distinct values		0
[integer]	min < med < max: 3 < 10.5 < 21 IQR (CV) : 6.8 (0.4)			(0%)
outcome	Mean (sd) : 7.9 (6.7)	18 distinct values		0
[integer]	min < med < max: 0 < 7 < 23 IQR (CV) : 10.8 (0.8)			(0%)

## Exploratory Data Analysis

Through Table 2, we can know the mean of the number of leprosy bacilli before treatment for the Rifampin group is 9.3 units, and the standard error is 4.76 units; the mean of the number of leprosy bacilli after treatment for the Rifampin group is 5.3 units, and the standard error is 4.64 units. Hence, there is an obvious trend that the mean of the number of leprosy bacilli decreases 4 units from before to after treatment, and the standard error is similar from before to after the treatment.

The mean of the number of leprosy bacilli before treatment for the Dapsone group is 10 units, and the standard error is 5.25 units; the mean of the number of leprosy bacilli after treatment for the Dapsone group is 6.1 units, and the standard error is 6.15 units. Hence, there is an obvious trend that the mean of the number of leprosy bacilli decreases 3.9 units from before to after treatment, and the standard error increases from before to after the treatment.

The mean of the number of leprosy bacilli before treatment for the Placebo group is 12.9 units, and the standard error is 3.96 units; the mean of the number of leprosy bacilli after treatment for the Placebo group is 12.3 units, and the standard error is 7.15 units. Hence, the mean of the number of leprosy bacilli only decreases 0.6 units from before to after treatment, and the standard error increases from before to after the treatment.

After the observation of three groups, we can suppose that because treatments have different effects to different patients, there may be other reasons that affect the result of the experiment. On the other hand, through the experimental result of the patients in Placebo group, for the patients who didn't have therapeutic effect, the difference in the number of leprosy bacilli from before to after will be larger. As a result, the possibility of our suppose is high.

Table 2: The mean of the number of leprosy bacilli

	Rifampin	Dapsone	Placebo	Total
baseline	9.3(4.76)	10(5.25)	12.9(3.96)	10.73(4.79)
outcome	5.3(4.64)	6.1(6.15)	12.3(7.15)	7.9(6.67)

Through the Figure 1, we can know the number of leprosy bacilli in antibiotics (Rifampin or Dapsone) group is obviously decreasing, and don't have the trend of increasing. Otherwise, the distribution of Dapsone group after treatment is more disperse than Rifampin's. We suppose that the therapeutic effect was not fit to every patient, but most of patients got great therapeutic effect. For the Placebo group, the distribution becomes more disperse after few months. We think that for the patients who didn't get the treatment may have ability of self-healing, but they also may have probability of producing more leprosy bacilli.

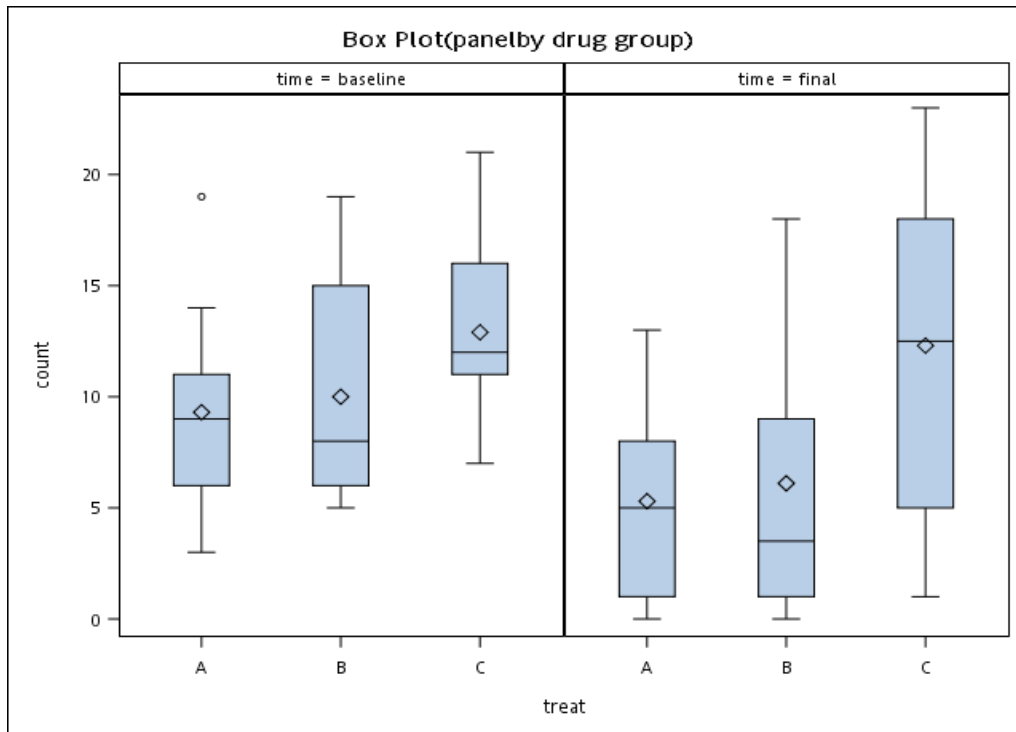


Figure 1: The number of leprosy bacilli before and after Box Plot

Through the Figure 2, the time that the number of leprosy bacilli decreases is similar for group A and group B, and we can also find that the degree of decrease in group C is smaller than group A and group B.

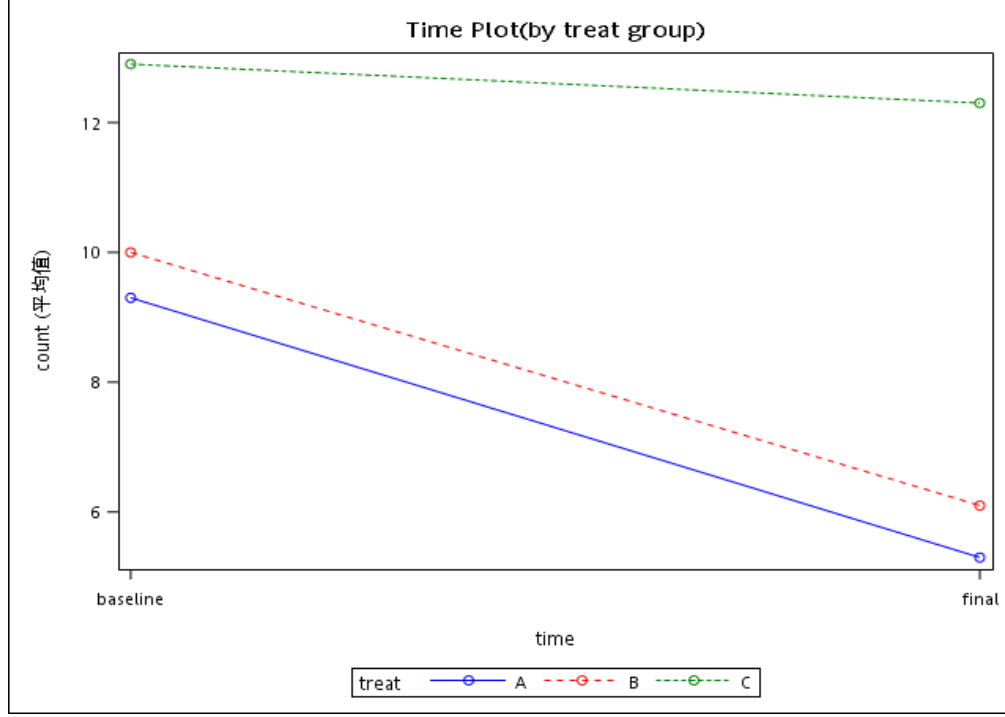


Figure 2: 治療前後痲瘋桿菌數量變動圖

## Method

### t-test

In order to observe whether there is any significant difference between antibiotics group and Placebo group, we have to use the value of variable, baseline, minus the variable, outcome, and get the reduce from before to after the treatment. Furthermore, we also use t-test to test the reduce from before to after the treatment for group A and group C, group B and group C, antibiotics group and Placebo group. For the below testing, we set significance level as 0.05:

Table 3: t-test for treatment groups

group	Variable	Variances	Method	ProbF	Probt
A	reduce	Equal	Folded F	0.139	0.106
B	reduce	Unequal	Satterthwaite	0.036	0.106
antibiotics	reduce	Unequal	Satterthwaite	0.017	0.093

The hypothesis testing for group A and group C:

$$\begin{cases} H_0: \sigma_A^2 = \sigma_C^2 \\ H_1: \sigma_A^2 \neq \sigma_C^2 \end{cases}$$

Through the result of t-test, we can know  $p\text{-value}=0.1394>0.05$ , so the variances for two groups are equal. Otherwise, we can use  $p\text{-value}=0.1056>0.05$  by Folded F method. As a result, we can infer that there is no significant difference of variance between group A and group C.

The hypothesis testing for group B and group C:

$$\begin{cases} H_0: \sigma_B^2 = \sigma_C^2 \\ H_1: \sigma_B^2 \neq \sigma_C^2 \end{cases}$$

Through the result of t-test, we can know  $p\text{-value}=0.0356<0.05$ , so the variances for two groups are unequal. Otherwise, we can use  $p\text{-value}=0.1056>0.05$  by Satterthwaite method. As a result, we can infer that there is no significant difference of variance between group B and group C.

The hypothesis testing for antibiotics group and Placebo group:

$$\begin{cases} H_0: \sigma_{\text{antibiotics}}^2 = \sigma_C^2 \\ H_1: \sigma_{\text{antibiotics}}^2 \neq \sigma_C^2 \end{cases}$$

Through the result of t-test, we can know  $p\text{-value}=0.0168<0.05$ , so the variances for two groups are unequal. Otherwise, we can use  $p\text{-value}=0.0927>0.05$  by Satterthwaite method. As a result, we can infer that there is no significant difference of variance between antibiotics group and Placebo group.

## Conclusion

Through t-test, we can realize that there is no significant difference for the reduce of the number of leprosy bacilli before and after treatment between antibiotics group (Rifampin group and Dapsone group) and Placebo group.

## B. Osteoporosis Clinical Trial

Researchers conducted a clinical trial to evaluate a new drug to treat osteoporosis in women past menopause. In a double-blind study, a group of women were assigned the treatment and a group of women were assigned the placebo. Both groups of women were provided with calcium supplements, given nutritional counseling, and encouraged to be physically active through the exercise programs made available to them. The study ran for three years, and the number of fractures occurring in each of those years was recorded. The length of each of the years, the corresponding risk periods, is 12 months. However, there were a few drop-outs in the third year, and those risk periods were set at 6 months. Data are in files CTosteoporosis3yrWide.csv and CTosteoporosis3yrLong.csv

## Introduction

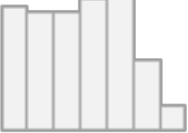
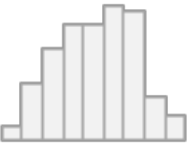





This is a trial that researches a new drug to treat osteoporosis in women past menopause, and the women are distributed to two groups, treatment group and Placebo

group. The researcher records the number of fractures every year for three years. Otherwise, because some patients drop out in the third year (7 people), some people's risk periods are set at 6 months. Through Table 4, the variables of the data include id, age, center, treatment, the number of fractures every year for three years, and there aren't any missing values in data. Furthermore, the number of people in each center and treatment group, and age are normally distributed. In each year, most of the number of fractures every year for three years is 0.

Table 4: Information of data

**Dimensions:** 214 x 7

**Duplicates:** 0

Variable	Stats / Values	Freqs (% of Valid)	Graph	Missing
id [integer]	Mean (sd) : 59.6 (34.6) min < med < max: 1 < 60.5 < 136 IQR (CV) : 57.5 (0.6)	124 distinct values		0 (0%)
age [integer]	Mean (sd) : 68.2 (9.4) min < med < max: 48 < 68 < 89 IQR (CV) : 15 (0.1)	38 distinct values		0 (0%)
center [character]	1. A 2. B	110 (51.4%) 104 (48.6%)		0 (0%)
treat [character]	1. p 2. t	102 (47.7%) 112 (52.3%)		0 (0%)
year1 [integer]	Mean (sd) : 0.1 (0.4) min < med < max: 0 < 0 < 2 IQR (CV) : 0 (3)	0 : 190 (88.8%) 1 : 19 ( 8.9%) 2 : 5 ( 2.3%)		0 (0%)
year2 [integer]	Mean (sd) : 0.1 (0.3) min < med < max: 0 < 0 < 2 IQR (CV) : 0 (4.1)	0 : 201 (93.9%) 1 : 11 ( 5.1%) 2 : 2 ( 0.9%)		0 (0%)
year3 [integer]	Mean (sd) : 0.1 (0.4) min < med < max: 0 < 0 < 2 IQR (CV) : 0 (3.5)	0 : 196 (91.6%) 1 : 14 ( 6.5%) 2 : 4 ( 1.9%)		0 (0%)

## Exploratory Data Analysis

Through Table 5, we can know there is no significant difference between ages, and the mean is 68 years old; the standard error is 9. Otherwise, for the patients in Placebo group, the mean of the number of fractures from after a year to three years is

decreasing, from 0.18 to 0.11, as the same as the standard error; for the patients in treatment group, the mean of the number of fractures from after a year to three years is decreasing significantly, from 0.1 to 0.03, as the same as the standard error, but those increase significantly.

Table 5: The statistical table for different ages and number of fractures

	Placebo	Treatment	Total
age	68.33(8.9)	68.04(9.91)	68.18(9.42)
year1	0.18(0.48)	0.1(0.33)	0.14(0.41)
year2	0.12(0.38)	0.03(0.16)	0.07(0.29)
year3	0.11(0.37)	0.1(0.35)	0.1(0.36)

Through Table 6, there is no significant difference between the number of people in center A and center B; there is also no significant difference between the number of people for each treatment in two centers.

Table 6: The statistical summary of the people in each group and center

center/treat	p	t	Total
Number of A	162.0	168.0	330
Percentage of A	25.2	26.2	51.4
Row Percentage of A	49.1	50.9	.
Column Percentage of A	52.9	50.0	.
Number of B	144.0	168.0	312
Percentage of B	22.4	26.2	48.6
Row Percentage of B	46.1	53.9	.
Column Percentage of B	47.1	50.0	.
Total	306.0	336.0	642
Total Percentage	47.7	52.3	100

Through Table 7, we can find that in Placebo group, as you take medicine for more longer year, the number of people who had fractures for one or two times at first will decrease, and the number of people who had fractures for zero times at first will increase. Hence, we can realize that the number of people who had fractures has a trend of decrease. On the other hand, for the patients in treatment group, although the number of people who had fractures for zero times increases, at the third year, it returns back to the value, which is similar to the first year.

Table 7: The statistical summary every year for each group

treat	year/count	0	1	2
P	1	88	10	4

.	2	92	8	2
.	3	93	7	2
t	1	102	9	1
.	2	109	3	0
.	3	103	7	2

Through Figure 3, for the both treatment and Placebo group, we can find that the number of fractures for patients who took the medicine will obviously decrease at the second year. However, at the third year, only the number of fractures in Placebo group still decrease; the number of fractures in treatment group has a trend that increase to the number at the first year.

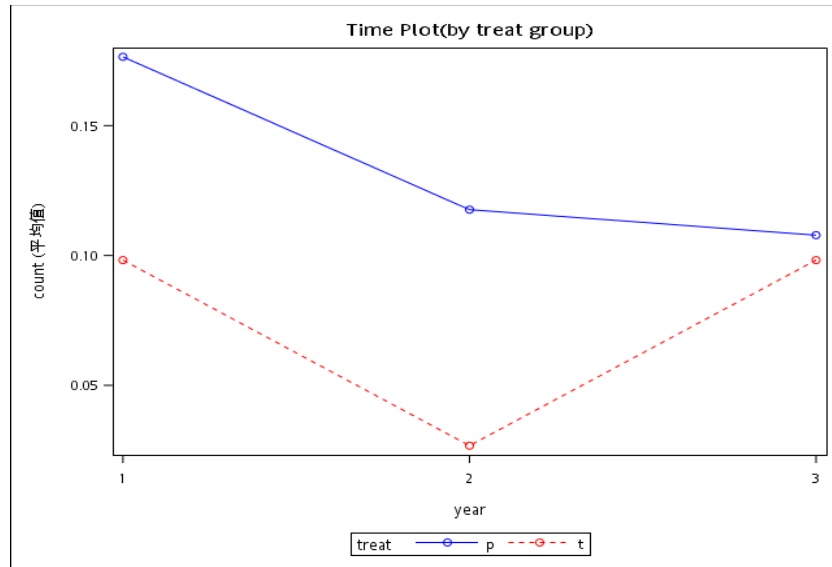


Figure 3: The change of the number of fractures for each year

Through Figure 4, the mean of the number of fractures for the patients in treatment group at center A and center B obviously decreases, but it increases at the third year. Furthermore, the value at center A at the third year even exceed the value at the first year. Otherwise, for the patients in Placebo group, the mean of the number of fractures decreases. However, it has little change in center A from year 2 to year 3, and that in center B is still decreasing.



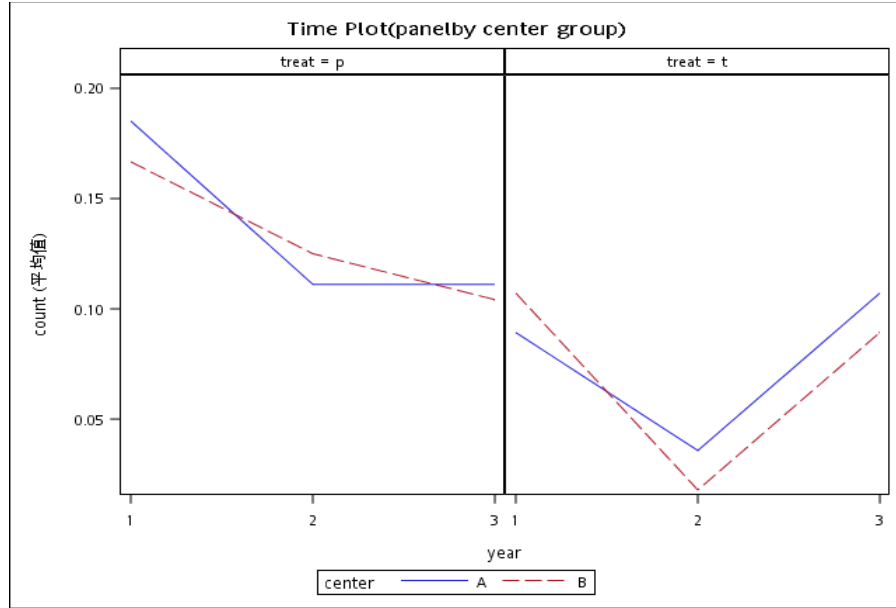


Figure 4: the change of the number of fractures at each center

## Method

### Wald test

We use fixed effect model to process Wald test in this research. We assume that the model uses the Poisson distribution and exponential random variables; the center, group, and the number of fractures are categorical variables; the time is continuous variable. Through Table 8, we can find that there only exists significant difference between groups, and the mean of number of fractures in Placebo group is obviously higher than it in treatment group. Furthermore, other variables to the mean of the number of fractures don't exist significant difference.

Table 8: Summary of the model

Coefficient	Category	Estimate	Standard Error	95% Confidence Interval		Z	Pr >  Z
				Lower Bound	Upper Bound		
Intercept	.	-4.109	1.082	-6.229	-1.989	-3.8	0.0001
age	.	0.021	0.015	-0.008	0.050	1.44	0.150
center	A	0.066	0.300	-0.522	0.653	0.22	0.826
center	B	0.000	0.000	0.000	0.000	.	.
treat	p	0.600	0.291	0.030	1.170	2.06	0.039
treat	t	0.000	0.000	0.000	0.000	.	.
year	1	0.276	0.296	-0.304	0.857	0.93	0.350
year	2	-0.383	0.398	-1.163	0.397	-0.96	0.335
year	3	0.000	0.000	0.000	0.000	.	.

## Conclusion

Because we don't know the situation at the baseline, although we can infer there is significant difference between groups through the model, we still can't infer whether that treatment can treat the osteoporosis. If we assume that there is no significant difference for the number of fractures of the whole patients at baseline, we can infer that the treatment can treat the osteoporosis successfully.